



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2011

Zum Mehrwert in der Ton-Bild-Beziehung

Flückiger, Barbara

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-53139>

Book Section

Published Version

Originally published at:

Flückiger, Barbara (2011). Zum Mehrwert in der Ton-Bild-Beziehung. In: Butte, M; Brandt, S. Bild und Stimme. Paderborn: Wilhelm Fink, 161-177.

Zum *Mehrwert* in der Ton/Bild-Beziehung

Barbara Flückiger

Ohne Zweifel leben wir in einer visuell dominierten Kultur. Das hat dazu geführt, dass während Jahrzehnten audio-visuelle Repräsentationsformen – vor allem Filme – primär als Bilder und Bildsequenzen diskutiert wurden, während man die Tonspur vernachlässigte. Wenn man die Klangelemente überhaupt untersuchte, fokussierten die Texte auf die Abhängigkeit der akustischen Informationen von der visuellen Repräsentation und allenfalls von der narrativen Konstruktion. Es dominierte also eine hierarchische Betrachtungsweise. Einen wichtigen Impuls zur Verlagerung der Diskussion leistete Michel Chion, hauptsächlich mit *L'Audio-vision*¹, das bereits 1994 in englischer Übersetzung² vorlag und daher auch im angelsächsischen Raum rezipiert wurde:

Der Film, das Fernsehen und die audiovisuellen Medien im Allgemeinen wenden sich nicht nur ans Auge. Sie wecken im Zuschauer – dem ›Audio-Zuschauer‹ – eine spezifische Wahrnehmungsweise, die ich in diesem Buch Audio-vision nennen werde. [...] Der Gegenstand dieses Buchs ist es zu zeigen, wie im audiovisuellen Zusammenspiel eine Wahrnehmungsmodalität die andere beeinflusst und verändert:

Man ›sieht‹ nicht mehr dasselbe, wenn man es gleichzeitig hört und man ›hört‹ nicht mehr dasselbe, wenn man es gleichzeitig sieht.³

Bis heute sind einige Basiskonzepte gültig, die Chion eingeführt hatte, namentlich die *Synchrese* (*synchrèse*) und der *Mehrwert* (*valeur ajoutée*). Sie tragen dazu bei, die Wahrnehmung audio-visueller Geflechte differenzierter zu untersuchen. Besonders ermöglichen sie es, die Interaktion von visuellen und auditiven *Cues*, also Hinweisen, in der spezifischen, mehrmodalen Anordnung in Filmwerken und anderen Medien zu untersuchen und zu verstehen. Allerdings lässt sich die Analyse dieser Interaktion auf der Basis von wahrnehmungspsychologischen Einsichten und neueren semantischen Konzepten weiter entwickeln.⁴ Zusätzlich sind technische Bedingungen einzubeziehen, denn mit der Ausdifferenzierung der Verfahren und des Dispositivs mit Stereo-Surround-Sound haben sich die Interaktionsmöglichkeiten signifikant verändert.⁵ Außerdem ist der kulturelle Rahmen zu berücksichtigen, in dem die Ton/Bild-Relation erscheint, weil dieser Rahmen selbst und insbesondere die kulturell kodierten Repräsentationsformen, die er hervorgebracht hat, maßgeblich zur Rezeptionstätigkeit des Zuschauers beitragen. Und schließlich ist die Frage zu stellen, ob jegliche Ton/Bild-Kombination möglich sei, oder wann die optischen und akustischen Repräsentationen in ihre Bestandteile zerfallen, weil der Zuschauer sie nicht mehr in einen (narrativen) Zusammenhang bringen kann; hier entsteht eine andere Ordnung der Bild/Ton-Relation – jenseits der kohärenten Zuschreibung.

Die Frage nach der audio-visuellen Verknüpfung ist von einiger Brisanz, denn in der Praxis, besonders des US-amerikanischen Mainstreamfilms, überwiegt die *Substitutionstechnik*, deren bekanntestes Beispiel wohl die Kokosnussschalen sind, mit denen das Geräusch von Pferdehufen nachgeahmt wird. Mit anderen Worten: Die Geräusche, die wir im Kino hören, stammen selten von den Objekten, die wir sehen, sondern werden durch akustisch *ähnliche* Geräusche ersetzt. Wie verändert sich diese *Ähnlichkeit* im Laufe der technischen und ästhetischen Entwicklung des Erzählkinos?

***Synchrese* und *Mehrwert*. Wahrnehmungspsychologische und semantische Aspekte**

Werden akustische und visuelle Reize präsentiert, sind zwei Phänomene zu beobachten, die Chion mit den Begriffen *Synchrese* und *Mehrwert* fasst:

Die *Synchrese* (ein Wort, das ich aus den Begriffen Synchronismus und Synthese zusammengesetzt habe) ist die unwiderstehliche und spontane Verbindung, die zwischen einem akustischen und einem kurzen optischen Phänomen entsteht, wenn beide zeitgleich auftreten, und zwar unabhängig von jeder rationalen Logik.⁶

Als *Mehrwert* bezeichne ich hier den Ausdrucks- und Informationswert, mit dem ein Ton ein gegebenes Bild auf verschiedenen Ebenen *anreichert*. Das kann soweit gehen, dass man über den unmittelbaren Eindruck oder die Erinnerung annimmt, dass diese Information oder dieser Ausdruck sich »natürlich« aus dem ergebe, was man sieht, und schon im Bild allein enthalten sei. Dadurch kann sogar der offensichtlich falsche Eindruck erweckt werden, dass der Ton überflüssig sei und eine Bedeutung nur verdopple, die er in Wirklichkeit herbeiführt oder schafft, sei es von Grund auf oder über die Differenz zum Sichtbaren. Dieses Phänomen des *Mehrwerts* funktioniert vor allem bei synchroner Darbietung von Bild und Ton über das Prinzip der *Synchrese*, die es erlaubt, eine unmittelbare und notwendige Beziehung zu schaffen zwischen dem, was man sieht, und dem, was man hört.⁷

Gemäß Chion sind also zwei Prinzipien am Werk, die einerseits auf der Wahrnehmungsebene anzusiedeln sind – nämlich die zeitliche Koinzidenz von optischem und akustischem Reizangebot – und andererseits aus semantischer Sicht als Bedeutungsmodifikation zu begreifen sind, weil Ton und Bild einander gegenseitig beeinflussen.

Aus wahrnehmungspsychologischer Sicht aufschlussreich ist ein Experiment, der so genannte *McGurk-Effekt*, den Harry McGurk und John MacDonald 1976 beschrieben haben und der als eine sehr sinnfällige Illustration der Verflechtung von Hören und Sehen funktioniert.⁸ In diesem Experiment untersuchten sie die Ton/Bild-Interaktion anhand der gesprochenen Sprache, und zwar indem sie eine Tonaufnahme eines Sprechers, der »Ba Ba Ba« sagte, mit dessen Bildaufnahme kombinierten, in welcher er »Ga Ga Ga« lautete. Die Mehrheit der Probanden kombinieren die Ton- und Bildinformation zu »Da Da Da«. Es kommt also in der intermodalen Assoziation⁹, das heißt der Verknüpfung von Sinnesdaten aus verschiedenen Wahrnehmungsbezirken, bei den Rezipienten zu

einer schlichten Modifikation der beiden Informationskanäle, die sich nur aus dem Zusammenspiel erklären lässt. Der McGurk-Effekt entspricht einer normalen Wahrnehmungssituation im Kino, in welcher die sichtbare oft eigentlich nicht der gehörten Quelle entspricht, da Geräusch und Bild getrennt voneinander aufgezeichnet werden.¹⁰ Ähnliches gilt für die Praxis der Nachsynchronisation von Dialogen.

Grundlagen zu heutigen Annahmen über die Wahrnehmungspsychologie der intermodalen Assoziation wurden bereits im 19. Jahrhundert gelegt. Eine Schlüsselfigur ist der Psychophysiker Hermann von Helmholtz, der 1862 eine grundlegende Fragmentierung der Wahrnehmung beschrieb.¹¹ Er bezeichnete die spezifischen Energiespektren, für die jeweils eine Sinnesmodalität empfänglich ist, als Qualitätenkreise, zwischen denen es weder einen kontinuierlichen Übergang, noch objektivierbare Vergleiche gebe, wie zum Beispiel zwischen blau und süß oder rund und laut.¹² Daraus folgt, dass die audio-visuelle Repräsentation sich aus wahrnehmungspsychologischer Sicht nicht grundsätzlich von der Alltagserfahrung unterscheidet. Vielmehr ist es so, dass sich das audiovisuelle Dispositiv des Kinos historisch aus den Einsichten der Psychophysik im 19. Jahrhundert entwickelt hat, also die Reizvielfalt der physikalischen Welt bereits so filtert und aufbereitet, dass sie optimal auf die menschliche Wahrnehmungstätigkeit zugeschnitten ist. Oder anders ausgedrückt ist hier nicht ein kulturell oder technologisch determiniertes Konzept zu beobachten, sondern vielmehr war die Entwicklung zuallererst biologisch und psychophysisch gestützt.

Noch radikaler als das Konzept der Qualitätenkreise ist das Prinzip der undifferenzierten Kodierung, das – vor Helmholtz – der Biologe Johannes Müller¹³ bereits in den 30er Jahren des 19. Jahrhunderts formuliert hatte:

Die Erregungszustände aller Rezeptoren codieren nur die Intensität, aber nicht die physikalische oder chemische Natur der Erregungssache: Codiert wird nur ›So-und-so-viel‹ an dieser Stelle meines Körpers, aber nicht ›Was‹.¹⁴

Sinnesempfindungen sind dieser konstruktivistischen Auffassung gemäß nicht irgendwie ›natürlich‹ in der Umwelt vorhanden, sondern entstehen nur im Wahrnehmungssystem. Schon Sir Isaac Newton hatte im 17. Jahrhundert festgestellt, dass Lichtstrahlen nicht farbig sind, sondern dass Farbe ausschließlich als Phänomen der Wahrnehmung begriffen werden muss.

Wenn also die Reize so spezifisch und fragmentiert verarbeitet werden, stellt sich die Frage, wie das Wahrnehmungssystem die Eindrücke aus den unterschiedlichen Sinnesbezirken miteinander verknüpft. Tatsächlich werden auditive und visuelle Stimuli von neuronalen Prozessen in einen Zusammenhang gebracht. Auf welcher Stufe der neuronalen Verarbeitung dies geschieht, ist Thema verschiedener Debatten. In einem Vortrag, den Steffen Lepa und Christian Floto an der Tagung der Gesellschaft für Medienwissenschaft 2003 präsentierten¹⁵, postulierten sie mit Verweis auf eine Studie an Katzen eine frühe Integration auf der Stufe des *Superior Colliculus*. Allerdings ist dort gemäß *Neuroscience*¹⁶, dem Standardwerk der Neurowissenschaft, lediglich die instinktive Orientierungsreaktion angesiedelt, nicht aber die Objektidentifikation, die eigentlich für die audiovisuelle Wahrnehmung von Medien entscheidend ist. Purves et al. beschreiben vielmehr verschiedene Assoziationskortexe, in denen Potentiale aus verschiedenen primären und sekundären Kortexen prozessiert werden.¹⁷ Es ist in diesem Zusammenhang aber zu betonen, dass selbst in der avanciertesten Neurowissenschaft solche komplexen Verarbeitungsmuster wegen grundlegender methodischer Probleme nicht restlos geklärt sind, und wir als GeisteswissenschaftlerInnen uns dieser Erkenntnisse nur vorsichtig bedienen sollten. Jedenfalls scheint es im hier diskutierten Zusammenhang eine wichtige Erkenntnis zu sein, dass man sich die kortikalen Strukturen als plastische Konnektionen vorstellt – im Unterschied zu instinktiven Verhaltensmustern auf frühen Verarbeitungsstufen, die fest verdrahtet sind (*hard-wired*). Diese plastischen Konnektionen werden auf der Basis von Lernprozessen gebildet. Der Neurologe Wolf Singer nimmt an,

dass räumlich verteilte Nervenzellen ihre Antworten synchronisieren müssen, wenn sie sich an der Kodierung einer kohärenten Figur beteiligen. [...] Der Grund ist, dass gleichzeitig eintreffende synaptische Potentiale sich in nachgeschalteten Zellen besonders gut summieren.¹⁸

Das wiederholt auftretende synchrone Feuern von neuronalen Verbänden führt ebenfalls zu einer verstärkten Gewichtung dieser Konnektionen und damit zu einer immer schneller und automatischer werdenden Informationsverarbeitung von bereits gefestigten Mustern. »Wissen über konsistente Beziehungen zwischen Phänomenen in der Welt wird auf diese Weise internalisiert und steht hinfort für die Interpretation sensorischer Signale zur Verfügung«.¹⁹

Gleichzeitig sind auch inhibitorische Verbindungen aktiv, um Merkmale, die sich ausschließen, voneinander zu trennen.

Grundlage der intermodalen Assoziation ist die zeitliche Übereinstimmung des Auftretens von akustischen und optischen Stimuli, wie es im Konzept der *Synchrese* beschrieben wird. Allerdings gibt es verschiedene Hinweise darauf, dass bei dieser Übereinstimmung eine gewisse Flexibilität möglich ist. So führen Lepa und Floto²⁰ verschiedene Experimente an, die die Idee eines gewissen Zeitfensters empirisch stützen, das Ernst Pöppel als zeitliche Toleranzgrenze für eine Verknüpfung in der Rezeption bei etwa 30 ms ansetzt.²¹ Eine solche Toleranz wäre bei einer Aufnahme Frequenz von 24 B/s weniger als ein Filmbild; die Erfahrung lehrt jedoch, dass ungeübte Rezipienten eine Asynchronität zwischen Bild und Ton sogar erst ab einer Differenz von zwei bis drei Bildern wahrnehmen. Durch wiederholtes Auftreten von ähnlichen oder gleichen Reizkonfigurationen – zunächst in der aktiven Exploration der Umwelt, dann aber gleichermaßen in der Rezeption von medialen Repräsentationen – entwickeln sich mehr oder weniger feste Strukturen im Langzeitgedächtnis, welche die Erwartungshaltung der Zuschauer, wie die optischen und akustischen Reize zusammenpassen, konditionieren.

Semantisch kann der *Mehrwert* im Sinne Chions als eine dynamische Form der Bedeutungserzeugung verstanden werden. Das heißt, es findet ein energetischer Fluss zwischen den beiden Konzepten – einem optischen und einem akustischen – statt, der zu etwas Drittem führt, das sich weder aus dem Einen noch aus dem Anderen erklären lässt. Es ist dies ein Prozess der gegenseitigen Modifikation, wie ihn beispielsweise Max Black in seiner interaktionistischen Metapherntheorie beschreibt.²² Durch eine Interaktion werden bestimmte Merkmale des Objekts hervorgehoben – beispielsweise dessen Materialität –, während andere unterdrückt werden. Wo aber liegen die Grenzen der Interaktion? Oder anders gefragt, ist jedwede Ton/Bild-Kombination möglich?

In Andrej Tarkowskis Film *Solaris* (UdSSR 1972) sind die Bewegungen einer Frau mit Glasgeräuschen synchronisiert. So entsteht der schwer nachvollziehbare Eindruck, dass der Körper der Frau aus Glas bestehe, obwohl äußerlich keine Veränderung der Materialität zu erkennen ist. Eine solch experimentelle Anordnung formt sicherlich einem Grenzwert dessen, was möglich ist, ohne dass die Verbindung zwischen Ton und Bild zusammenbricht. Allerdings liefert die filmische Narration von *Solaris* eine Erklärung für diese

ungewöhnliche Transformation, denn die Frau ist eigentlich bereits gestorben und ihr *Bild* entspricht eher der Projektion ihres Mannes. Nur dieser Erklärung ist es zu verdanken, dass die ZuschauerInnen die klangliche Veränderung als plausibel akzeptieren. Von anderer Art sind die Ton/Bild-Assoziationen, die sich im Animationsfilm etabliert haben, wo sowohl die stilisierten Bilder als auch die Töne sich deutlich von jedweder Alltagserfahrung unterscheiden und fern von jedem Realitätseffekt eine ungeheure kreative Freiheit ermöglichen, die sich weder den physikalischen Gesetzen zu beugen, noch der lebensweltlichen Realität unterzuordnen hat. Ein Sonderfall schließlich sind Subjektivierungen, also die Verlagerung der Erzählperspektive in die Wahrnehmungsperspektive einer Filmfigur. Eine übergeordnete Strategie zur Subjektivierung ist die Dissoziation von Ton und Bild²³, ein deutlich wahrnehmbares Auseinanderdriften von akustischer und optischer Repräsentation. Diese Dissoziation kann auf gelernten Mustern wie Hall beruhen, der ohne eine sichtbare räumliche Transformation hinzugemischt wird, oder sie kann subjektive Erfahrungen bei extremem Stress, in psychischen Grenzerfahrungen, Träumen oder bei Drogenkonsum imitieren.

Auf der anderen Seite des Spektrums – also der maximalen Transformation diametral entgegengesetzt – ist die Redundanz angesiedelt, die scheinbar keinen *Mehrwert* aufweist, eine Möglichkeit, die Michel Chion bezweifeln würde.²⁴ Um das Konzept der Redundanz zu diskutieren, ist es sinnvoll, die verschiedenen Schichten eines Klangobjekts²⁵ zu analysieren, die sich aus semantischen und sensorischen Aspekten zusammensetzen. Als Ausgangspunkt hierfür kann die Sprache dienen, bei welcher sich die Bedeutung – also der Inhalt oder die Botschaft – sehr leicht von den als Prosodie bezeichneten klanglichen Elementen unterscheiden lässt. Die klangliche Ebene kann Informationen über den emotionalen Zustand des Sprechers oder der Sprecherin und über seinen oder ihren sozialen oder ethnischen Hintergrund vermitteln. Zusätzlich ist der Klang der Stimme unmittelbar an die körperliche Kondition gebunden und ein weitgehend verlässlicher Indikator des Geschlechts, des Alters und möglicherweise auch des gesundheitlichen Zustands der sprechenden Person. In ähnlicher Weise lassen sich semantische und sensorische Informationen von Klangobjekten unterscheiden. Eine sich schließende Tür kann bedeuten, dass jemand kommt oder jemand eingeschlossen wird. Wie ich in einer Fallstudie zum Geräusch von Türen in *The Silence of the Lambs* (Jonathan Demme, USA 1991) gezeigt habe, können Türgeräusche darüber hinaus eine

symbolische Bedeutung tragen, die von ihrer gesellschaftlichen Funktion des Überschreitens von Grenzen zwischen sozialen Räumen oder des Wegschließens von gesellschaftlich geächteten und unerwünschten Personen stammen kann.²⁶ Zusätzlich vermitteln Türgeräusche auf sensorischer Ebene Informationen über die materielle Beschaffenheit, die Größe, das Gewicht und den Zustand. Quiet-schende Türen sind genregebundene Stereotypen und als intertextuell etablierte Zeichen höherer Ordnung zu begreifen: Sie rufen daher weitgehend zuverlässig ein bestimmtes Spektrum von Assoziationen und möglicherweise emotionalen Reaktionen des Zuschauers auf. In meiner Studie zum *Sound Design* habe ich deshalb folgende These formuliert: Je ausdifferenzierter ein Klangobjekt sowohl semantisch als auch sensorisch ist, desto größer ist der resultierende *Mehrwert* in der Interaktion mit der bildlich repräsentierten hypothetischen Quelle.²⁷ Deshalb hat ein sehr standardisiertes Klangobjekt mit geringer Auflösung nur wenig Einfluss auf die Objektwahrnehmung, weil es über eine Deixis im einfachen Sinne des »Da ist ein ...« nicht hinaus geht und weil auch die sensorischen Eigenschaften unspezifisch sind. Einen solchen Fall würde ich als Redundanz bezeichnen. Nehmen wir als Beispiel das berühmt-berüchtigte Klappern von Pferdehufen, vom Geräuschemacher mit Kokosnussschalen vertont: Von den vielen möglichen Varianten, die Pferdehufe je nach Situation und Beschaffenheit des Bodens erzeugen können, repräsentiert es einen faden Mittelwert. Üblicherweise ist das abgebildete Pferd in einem klassischen Hollywoodfilm genauso typisiert. Daher überlappen sich akustische und optische Repräsentation weitgehend, obwohl sie selbstverständlich unterschiedliche Sinnesmodalitäten ansprechen. Oder anders ausgedrückt: Redundanz meint, dass Bild und Klangobjekt ein *inneres Bild* – ein Schema – nicht anreichern, sondern lediglich verdoppeln; deshalb findet auch keine wechselseitige Transformation statt und es entsteht kein signifikanter *Mehrwert* im Sinne einer semantisch-expressiven oder informativen Anreicherung.

Technische Grundlagen und Klangsubstitution als Praxis

Jedenfalls scheint die klangliche Ausdifferenzierung von Geräuschen ein historischer Prozess zu sein, der sowohl von technischen Limitierungen als auch von kulturellen Faktoren bestimmt ist.

Wie Chion darlegt, ist selbst unter Film- und MedienwissenschaftlerInnen die Ansicht verbreitet, dass Töne und ganz besonders Geräusche irgendwie »natürlich« da seien und wie selbstverständlich

vom filmischen Raum, den Objekten und Lebewesen auf der Leinwand ›ausgingen‹. Tatsächlich gibt es eine lange Tradition, Geräusche zu substituieren, eine Tradition, die bis ins antike Theater der Griechen zurück reicht. Schon im 1. Jahrhundert v. Chr. hatte Heron von Alexandrien hydraulische Maschinen entwickelt, um Windgeräusche, Donner oder Vogelgezwitscher zu erzeugen.²⁸ Im Lauf der Zeit wurde diese Tradition durch weitere Instrumente und Techniken verfeinert. Zwar waren Geräusche und Toneffekte im Kino der Stummfilmzeit mit großer Wahrscheinlichkeit weitgehend abwesend, aber wie ein rares Dokument von 1914 – nämlich S. de Serks *Les Bruits de coulisses au cinéma* – belegt, gab es schon zu dieser Zeit ein überraschend reichhaltiges Arsenal künstlich hergestellter Geräusche, die Geräuschemacher (*foley artists*) in späteren Dekaden aufgegriffen und weiter entwickelt haben.²⁹ Mit der Einführung des Tonfilms in den späteren 1920er Jahren schien zunächst eine Praxis der vollständigen Originalaufnahme vorzuherrschen. So ließ Alfred Hitchcock für *Blackmail* (GB 1929) Mikrofone vor dem Studiogelände aufstellen, um die Strassengeräusche live aufzunehmen. Wegen technischer Limitierungen sind diese frühen Tonspuren von einer deutlichen Komplexitätsreduktion geprägt. Der frühe Mono-Lichtton konnte nur einen beschränkten Frequenzumfang und eine stark komprimierte Dynamik wiedergeben. Wegen ihrer nicht-harmonischen, unstrukturierten Gestalt litten die Geräusche besonders unter diesen Einschränkungen, denn deren Identifikation ist unmittelbar an eine hohe klangliche Auflösung gebunden. Überdies führt die Anordnung aller Klangelemente auf einer Spur im Monoverfahren dazu, dass sie sich gegenseitig überlagern und somit maskieren. Favorisiert wurden deshalb kulturell gut verankerte Klangobjekte mit einfacher akustischer Gestalt, besonders Signallaute wie Autohupen und Telefonklingeln oder Tierstimmen. Außerdem haben einige Jahre später – in den frühen 1930er Jahren – die Praktiker festgestellt, dass Originaltonaufnahmen nicht zwingend die überzeugendsten Resultate liefern: »Wir geben es ungern öffentlich zu, aber es scheint so zu sein, dass Filmgeräusche am natürlichsten klingen, wenn sie das Ergebnis von intelligenten und durchdachten Kunstgriffen sind.«³⁰

Weil das System mit so vielen Beschränkungen zu kämpfen hatte, war Redundanz beziehungsweise deutliche Referenz aus informationstheoretischer Sicht sinnvoll: Nach einer kurzen, experimentellen Übergangsphase – der französische Filmtheoretiker Laurent Jullier nannte sie *Interregnum*³¹ – fügte sich die Tonarbeit in das bereits etablierte Continuity-System des klassischen Hollywood ein,

dessen Ziel es ist, die Erzählinstanz zu maskieren und die erzählerischen Elemente so anzuordnen, dass sie wie von selbst ein kohärentes Ganzes bilden und die fragmentierte und addierte Darstellungsweise des Films verschwinden lassen.³² Dieses Bestreben kann als Naturalisierungstendenz zur Schaffung einer hermetischen, kontinuierlichen *Diegese*, das heißt eines geschlossen wirkenden Raum-Zeit-Kontinuums beschrieben werden. In diesem System wurden die Klangelemente strikt auf die erzählerischen Ziele hin funktionalisiert und der hierarchischen Ordnung zur Lenkung der Zuschaueraktivität unterworfen. Aus diesen Zielsetzungen ergab sich die Regel *See a dog – hear a dog*, die eine reziproke Beziehung zwischen Ton und Bild beschreibt: Kein Hund kommt ins Bild ohne nicht wenigstens einmal kurz zu knurren, jedes Pferd wiehert oder schnaubt, jeder Hahn kräht. Mit Ausnahme weniger gut etablierter Klangobjekte ist im Erzählkino darüber hinaus auch kein Geräusch *akusmatisch* – also ohne visuelle Verankerung seiner Quelle – zu hören.

Sound Design und das souveräne Klangobjekt

Chion hat den neueren Mainstreamfilm, den er nicht sonderlich zu schätzen scheint, weitgehend aus seinen Untersuchungen ausgeblendet – zu Unrecht, wie ich meine. Denn wie ich – ausgehend von der Analyse eines Korpus von 94 Filmen – in meiner Untersuchung *Sound Design* zeige, hat in der Praxis ab Mitte der 1970er Jahre ein deutlicher und aufschlussreicher Wandel eingesetzt, weil sich das Sound Design³³ als eigenständige Domäne etablierte. Dieser Wandel ist einer Reihe von kulturellen Einflüssen zuzuschreiben, so beispielsweise der Rockmusik oder dem europäischen Art-House-Kino wie etwa der französischen Nouvelle Vague. Weitere Faktoren für den Wandel dürften zum einen die sich verändernde Arbeitsorganisation mit flacherer Hierarchie gewesen sein, die dazu führte, dass die Sound Designer als kreative Partner in den Prozess einbezogen wurden und nicht als einfache Erfüllungsgehilfen eines standardisierten Systems zu agieren hatten, zum anderen die zunehmende Medienkompetenz des Publikums, das zunehmend bereit war, auch komplexere Reiz- und Strukturangebote zu entschlüsseln. Im technischen Bereich führte die Aufrüstung der Consumer-Elektronik mit dem HiFi-Standard ab den späten 1950er Jahren zu veränderten Hörgewohnheiten des Publikums und übte einen Innovationsdruck auf die Filmindustrie aus. Aufschlussreich ist jedoch, dass mit den Magnetton-Mehrkanal-Verfahren, die bereits in den 1950er Jahren mit den Breitwandformaten wie CinemaScope u.a. Verbreitung gefunden

hatten, kein signifikanter Wandel des tonästhetischen Repertoires einherging, obwohl diese Formate über eine deutlich bessere klangliche Auflösung verfügten und komplexere Anordnungen durchaus erlaubten. Das legt den Schluss nahe, dass die kulturellen Faktoren bedeutsamer waren als die technischen.

Das *souveräne Klangobjekt*, das ich als abschließende Erweiterung der Bild/Ton-Interaktion betrachten möchte, und das im Umfeld des veränderten Tonbewusstseins in den 1970er Jahren entstand, war wilder und rauer und entzog sich der rigiden und stereotypen Stilisierung der früheren Dekaden.³⁴ Es war vielschichtig und produzierte einen sensorischen Überfluss, der sich der erzählerischen Funktionalisierung durchaus verweigert und mehr auf die unbewusste, emotional besetzte Schicht der Rezeption abzielt. Das viel zitierte Laserschwert aus der *Star Wars*-Saga ist nach wie vor eines der am besten einleuchtenden Beispiele, um die komplexe Ton/Bild-Interaktion zu diskutieren, die das *souveräne Klangobjekt* mit sich bringt.³⁵ Diese Diskussion wird es am Ende auch erlauben, den Bogen wieder zu schließen und die oben dargelegten wahrnehmungs- und neuropsychologischen Einsichten nochmals aufzugreifen. Im Film *Star Wars IV A New Hope* (George Lucas USA 1977) wird zunächst das unbekannte Objekt sprachlich benannt als »Your father's light sabre« und damit dem semantischen Register der Waffen zugeordnet. Allerdings besteht dabei eine deutlich wahrnehmbare Diskrepanz zwischen bildlicher Darstellung – das Objekt sieht wie eine Neonröhre aus – und behaupteter Funktion: Es fehlt das Merkmal »gefährlich«, das für Waffen obligatorisch ist. Über die Integration einer Reihe von Teilkonzepten und klanglichen Elementen wird das Objekt nach und nach audio-visuell ausdifferenziert und charakterisiert.

Der Sound Designer Ben Burtt, der dieses erste Laserschwert akustisch gestaltete, nahm als Grundelement mit einem Interferenzmikrofon³⁶ das Surren eines Projektors gemischt mit dem Brummen eines nicht funktionierenden Fernsehgeräts auf, indem er das Mikrofon synchron zu den Bewegungen im Bild durch das diffuse, inszenierte Klangfeld führte. Damit erzeugte er einen Dopplereffekt³⁷ mit minimalen klanglichen Veränderungen in exakter zeitlicher Anpassung an die visuelle Repräsentation, welche die Glaubwürdigkeit der reziproken Beziehung unterstützt. Zusätzlich entstand bei der Bewegung des Interferenzmikrofons im diffusen Klangfeld eine Phasing genannte Klangmodulation. Dieses *Phasing* ist normalerweise ein Charakteristikum von extrem schnell bewegten Objekten wie Düsenjets. Auf das Ein- und Ausschalten des Laserschwerts – bildlich



**1, 2 Mark Hamill
und Alec Guinness
in: George Lucas,
Star Wars. A New
Hope, 1977.**

dargestellt als Transformation des zunächst kleinen und unscheinbaren Metallstabs in die oben beschriebene Neonröhre – montierte er ein kurzes, unangenehmes Zischen. Außerdem ist immer, wenn der Lichtstrahl einen Feind berührt, ein ätzendes Knistern zu hören, ähnlich dem Geräusch, das entsteht, wenn durch Elektrizität oder große Hitze ein feuchter Gegenstand verschmort wird. Die Konsequenz dieser gewalttätig klingenden Berührungen wird in der ersten Anwendung des Laserschwerts auch bildlich konkretisiert in Form eines abgetrennten Arms und so eine kausal motivierte, narrative Folge hergestellt. [Abb. 1 und 2]

Die Situations- und vor allem die Objektkonstitution des phantastischen Gegenstandes vollziehen sich also über die Integration disparater Merkmale, die weit über die reine Hinweiskfunktion hinausgeht. Spezifische Merkmale etablieren Funktionen einschließlich ihrer Konsequenzen, Prozesse, einer hypothetischen materiellen Beschaffenheit und sinnlicher Qualitäten. Das Klangobjekt selbst besteht aus verschiedenen Teilkonzepten, die verschiedenen Prozessen zugeordnet werden: die Klangmodulation der Bewegung im Raum, das Zischen der Funktionsaufnahme und das Knistern der Berührung. Die klanglichen Qualitäten mobilisieren so auch situative und kausale Sinneserfahrungen aus der Alltagswelt: Die Assoziation von Geschwindigkeit und Potenz wird über das *Phasing* angeregt sowie die Assoziation von negativen Erfahrungen mit Strom und Hitze über das ätzende Knistern. Während die zeitliche Übereinstimmung von Ton und Bild die Kausalitätsbeziehung festigt, werden über das Zischen zwei verschiedene Zustände des Objekts – ein und aus – miteinander verbunden, das heißt Kontinuität geschaffen. Einmal etabliert, kann sich diese Transformation auch außerhalb des Bildfelds abspielen und wird dann nur noch akustisch repräsentiert. Diese asymmetrische Bedeutungszuordnung ist schon nach der ersten Exposition gefestigt. So wurde in *Star Wars IV A New Hope* die



hypothetische materielle Beschaffenheit des Laserschwerts über die Ähnlichkeit der abgebildeten Quelle mit der substituierenden Quelle (der Bildröhre des Fernsehers) erzeugt. Diese Übertragung kann als Assimilation – als Anpassung an bestehende mentale Repräsentationen von ähnlichen Objekten – verstanden werden. Nicht zuletzt ist das undefinierbare Zischen ein Signatur-sound des Science-Fiction-Films geworden und ruft deshalb über intertextuelle Verknüpfungen ein weiteres Feld von Assoziationen auf.

Sicherlich wird bei diesem Beispiel sofort deutlich, warum ein differenziertes vielschichtiges Klangobjekt einen höheren *Mehrwert* erzeugt als eine weitgehend standardisierte Verknüpfung. Ausgehend von den wahrnehmungspsychologischen Überlegungen können wir annehmen, dass der *Mehrwert* dem Aufwand entspricht, den der Rezipient zum Entschlüsseln der *polyphonie informationelle* (Roland Barthes) leisten muss. Etablierte Beziehungen wie die erwähnte Pferdehuf-Kokosnuss-Assoziation werden sofort und automatisch verarbeitet. Komplexere Verknüpfungen mit unscharfen oder mehrdeutigen Teilkonzepten auf beiden Seiten der audio-visuellen Repräsentation hingegen benötigen sehr viel mehr Aufmerksamkeit. Diese Beobachtung lässt sich mit dem Konzept der Verfremdung erklären, das der russische Formalist Viktor Šklovskij 1916 entwickelte.³⁸ Er postulierte in seinem Aufsatz *Die Kunst als Verfahren* am Beispiel der Dichtung, dass das Kunstwerk automatische und alltägliche Prozesse der Wahrnehmung unterbrechen und so die sinnliche Wahrnehmung irritieren und schärfen könne und solle.³⁹ Allerdings verblasst dieser Effekt, wenn sich neue Ausdrucksweisen etablieren und so neue Konventionen bilden, die wiederum stereotypisierte Formen hervorbringen.

So zeigte sich am Beispiel des *souveränen Klangobjekts* Lichtschwert, dass es zwar Grenzen der Bild/Ton-Interaktion des Erzählkinos gibt, die historisch, technisch und ästhetisch bedingt und auf Neuerungen und Innovation angewiesen sind. Diese Grenzen

sind aber nicht nur historisch gesehen elastisch, sondern sie werden auch von semio-pragmatischen Aspekten mitgeformt, die durch den kulturellen Rahmen des Films und dessen Tradition bestimmt sind. Die ästhetische Wahrnehmung der audiovisuellen Konfigurationen – besonders eines *souveränen Klangobjekts* wie dem Lichtschwert – ist in ähnlicher Weise wandel- und formbar, geprägt von Kohärenzen und Brüchen, Traditionen und Innovationen, Alltagserinnerungen und Neuerfahrungen, deren Verhältnis von Fall zu Fall betrachtet werden muss.

Endnoten

- 1 Michel Chion, *L'Audio-vision. Son et image au cinéma*, Paris 1990.
- 2 Ders., *Audio-vision. Sound on Screen*, übers. von Claudia Gorbman, New York 1994.
- 3 »Les films, la télévision et les médias audio-visuels en général ne s'adressent pas seulement à l'œil. Ils suscitent chez leur spectateur – leur ›audio-spectateur‹ – une attitude perceptive spécifique, que nous proposons, dans cet ouvrage, d'appeler l'audio-vision. [...] L'objet de ce livre est de montrer comment en réalité, dans la combinaison audio-visuelle, une perception influence l'autre et la transforme: on ne ›voit‹ pas la même chose quand on entend; on n' ›entend‹ pas la même chose quand on voit.« Chion, *L'Audio-vision* (Anm. 1), S. 3 [Übers. BF].
- 4 Basiskonzepte dieser Überlegungen gehen auf mein Forschungsprojekt zum Sound Design im amerikanischen Mainstreamfilm zurück: Barbara Flückiger, *Sound Design. Die virtuelle Klangwelt des Films*, Marburg 2001 [3. Auflage 2006].
- 5 Surround-Systeme verfügen neben dem Stereo-Signal über einen oder mehrere Surround-Kanäle, die über seitlich und hinten angeordnete Lautsprecher in den Kinosaal eingespielt werden. Vgl. Flückiger, *Sound Design* (Anm. 4), S. 50 ff.
- 6 »La synchrèse (mot que nous forgeons en combinant synchronisme et synthèse) est la soudure irrésistible et spontanée qui se produit entre un phénomène sonore et un phénomène visuel ponctuel lorsque ceux-ci tombent en même temps, cela indépendamment de toute logique rationnelle.« Chion, *L'Audio-vision* (wie Anm. 1), S. 55 [Übers. BF].
- 7 »Par *valeur ajoutée*, nous désignons la valeur expressive et informative dont un son enrichit une image donnée, jusqu'à donner à croire, dans l'impression immédiate qu'on en a ou le souvenir qu'on en garde, que cette information ou cette expression se dégage ›naturellement‹ de ce qu'on voit, et est déjà contenue dans l'image seule. Et jusqu'à procurer l'impression, éminemment injuste, que le son est inutile, et qu'il redouble un sens qu'en réalité il amène et crée, soit de toutes pièces, soit par sa différence même d'avec ce qu'on voit.« Chion, *L'Audio-vision* (Anm. 1), S. 8 [Übers. BF].
- 8 Harry McGurk und John MacDonald, *Hearing Lips and Seeing Voices*, in: *Nature* 246, 1976, S. 746–748. Ein Video des Effekts ist auf Youtube zu sehen: <http://www.youtube.com/watch?v=I1XWDOwH47Y> [31.10.2009].
- 9 Die inter- oder kreuzmodale Assoziation ist von der Synästhesie zu unterscheiden. Während in der intermodalen Assoziation die Reize aus unterschiedlichen Modalitäten verknüpft werden, ist die Synästhesie als Projektion eines Sinnesorgans in ein anderes zu begreifen, wie beispielsweise ein Farben-Hören. Dazu bedarf es einer spezifischen Begabung. Der Neurologe Richard E. Cytowic geht allerdings davon aus, dass die synästhetische Wahrnehmung beim Säugling zunächst die Regel ist und sich die spezifischen Areale zur Reizverarbeitung erst später ausdifferenzieren.
- 10 Siehe Abschnitt *Technische Grundlagen und Klangsubstitution als Praxis* im vorliegenden Artikel.
- 11 Hermann Ludwig Ferdinand von Helmholtz, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, Braunschweig 1862.
- 12 Eine Ausnahme von dieser Regel ist der tiefe Bassbereich akustischer Reize, der sowohl auditiv als auch taktil sowie zusätzlich über das vestibuläre System, das Gleichgewichtsorgan im Innenohr, wahrgenommen werden kann.
- 13 Johannes Müller, *Handbuch der Physiologie des Menschen für Vorlesungen*, Koblenz 1837–1840.
- 14 Heinz von Foerster, *Wahrnehmen*, in: Jean Baudrillard (Hg.), *Ars Electronica. Philosophien der neuen Technologie*, Linz 1989, S. 27–40, hier S. 35.
- 15 Steffen Lepa und Christian Floto, *Audio-Vision als Konstruktion. Grundzüge einer funktionalistischen Audioanalyse von Film und Multimedia*, Vortrag präsentiert an der Tagung der Gesellschaft für Medienwissenschaft am 11.9.2003.
- 16 Dale Purves, George J. Augustine, David Fitzpatrick, William C. Hall, Anthony-Samuel LaMantia, James O. McNamara und Leonard E. White (Hg.), *Neuroscience. Fourth Edition*, Sinauer Associates, Sunderland (Mass.) 2008, S. G-14.

Endnoten/Abbildungsnachweis

- 17 Ebd., S. 667.
- 18 Wolf Singer, Der Beobachter im Gehirn, in: Heinrich Meier und Detlev Ploog (Hg.), Der Mensch und sein Gehirn. Die Folgen der Evolution, München/Zürich 1997, S. 53.
- 19 Ebd., S. 61.
- 20 Lepa et al., Audio-Vision als Konstruktion (Anm. 15), S. 3.
- 21 Ernst Pöppel, Zeitlose Zeiten. Das Gehirn als paradoxe Zeitmaschine, in: Heinrich Meier und Detlev Ploog (Hg.), Der Mensch und sein Gehirn. Die Folgen der Evolution, München u.a. 1997, S. 67–98, hier S. 80 ff.
- 22 Max Black, Die Metapher [1954], in: Anselm Haverkamp (Hg.), Theorie der Metapher, Darmstadt 1996, S. 55–80.
- 23 Flückiger, Sound Design (Anm. 4), S. 395 ff.
- 24 Vortrag von Michel Chion an der Tagung *Utopia of Sound* in Wien am 28.5.2008.
- 25 Der Begriff Klangobjekt (*objet sonore*) geht auf Pierre Schaeffer zurück, den Begründer der *Musique concrète*: »Als Klangobjekt bezeichne ich hier den Ton selbst, seine akustische Beschaffenheit und nicht das materielle Objekt (ein Instrument oder eine andersgeartete Quelle), von dem es stammt.« [»Par objet sonore nous désignons ici le son lui-même, considéré dans sa nature sonore, et non pas l'objet matériel (instrument ou dispositif quelconque) dont il provient.«] Pierre Schaeffer, *Traité des objets musicaux. Essai interdisciplinaires*, Paris 1966, S. 23, 97 [Übers. BF].
- 26 Flückiger, Sound Design (Anm. 4), S. 309 ff.
- 27 Ebd., S. 143 ff.
- 28 Vgl. Reinhold Hammerstein, Macht und Klang. Tönende Automaten als Realität und Fiktion in der alten und mittelalterlichen Welt, Bern 1986.
- 29 Siehe <http://www.filmsound.org/foley> [Zugriff 18. März 2010]
- 30 »We perhaps dislike admitting it in public, but it seems to be a fact that the most natural appearing results obtained in this business are quite apt to be the result of intelligent and studied artifices.« Wesley C. Miller, *The Illusion of Reality in Sound Pictures*, in: Lester Cowan (Hg.), *Recording Sound for Motion Pictures*, New York 1931, S. 212 [Übers. BF].
- 31 Vgl. Jullier, Laurent, *Les Sons au cinéma et à la télévision. Précis d'analyse de la bande-son*, Paris 1995.
- 32 Zum Continuity-System siehe David Bordwell, Janet Staiger und Kristin Thompson, *The Classical Hollywood Cinema*, London 1985.
- 33 Sound Design umfasst die Erarbeitung eines tonästhetischen Gesamtkonzepts für die Bereiche *Sprache* und *Geräusch*, nicht aber für die Musik. Im Unterschied zum Leiter des Tondepartements im klassischen Studiosystem ist der Sound Designer nicht nur Manager, sondern auch zentrale kreative Instanz. Er entwickelt einen Stil, er schafft mit klanglichen Elementen dramaturgische Verbindungen zwischen Figuren, Orten und Objekten und erweitert die emotionale Dimension des Films mit subtilen Transformationen des Tonmaterials. Vgl. Flückiger, Sound Design (Anm. 4), S. 18.
- 34 Ebd., S. 121.
- 35 Vgl. im Folgenden die umfangreichere Fallstudie in: ebd., S. 144.
- 36 Ein Interferenzmikrofon hat eine stark gebündelte, keulenförmige Richtcharakteristik, erzeugt durch das Bauprinzip das den seitlich einfallenden Schall über Interferenzen dämpft.
- 37 Dopplereffekt: Veränderung der wahrgenommenen Tonhöhe einer sich schnell bewegenden Schallquelle
- 38 Viktor Šklovskij, Die Kunst als Verfahren [1916], in: Jurij Striedter (Hg.), *Russischer Formalismus. Texte zur allgemeinen Literaturtheorie und zur Theorie der Prosa*, München 1994, S. 5–35, bes. S. 11.

Abbildungsnachweis

- 1, 2 George Lucas, *Star Wars. A New Hope* 1977, Lucasfilm. Twentieth Century Fox, Screenshots.

